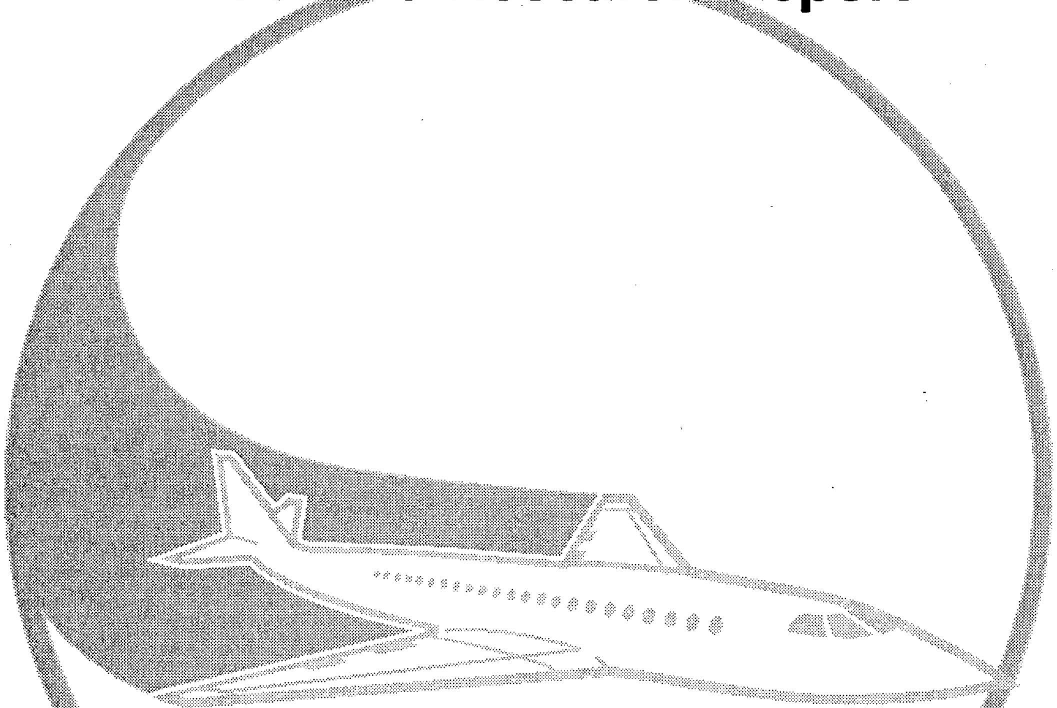
ote technical note techn

Plan View Display Baseline Research Report



September 1995

DOT/FAA/CT-TN95/45

U.S. Department of Transportation
Federal Aviation Administration
Technical Center
Atlantic City Airport, NJ 08405

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/CT-TN95/45		

4. Title and Subtitle

Plan View Display Baseline Research Report

5. Report Date

September 1995

6. Performing Organization Code

ACT-530

7. Author(s) J. Galushka, ACT-530, J. Frederick, ACT-230, Dr. R. Mogford, NYMA, Inc., and Dr. P. Krois, Crown Communications, Inc.

8. Performing Organization Report No.

DOT/FAA/CT-TN95/45

9. Performing Organization Name and Address

Federal Aviation Administration
Technical Center
Atlantic City International Airport, NJ 08405

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

F2202J

12. Sponsoring Agency Name and Address

Federal Aviation Administration Headquarters
AT Advanced Automation System Requirements
800 Independence Ave., S.W.
Washington, DC 20591

13. Type of Report and Period Covered

Technical Note
January 1995 - March 1995

14. Sponsoring Agency Code

ATR-320

15. Supplementary Notes

16. Abstract

The goal of this study was to provide baseline measurements of the current en route system consisting of the Host Computer System, the Plan View Display, and the M1 console. A set of measures were developed that related to six high level operational constructs: Safety, Capacity, Performance, Workload, Usability, and System Fidelity. In order to collect data on these measures, an air traffic control simulation was completed, using Washington Air Route Traffic Control Center (ZDC) airspace, controllers from ZDC, and a traffic scenario based on ZDC System and Analysis Recording (SAR) data.

Objective measures were reduced from SAR and Amecom tapes. Subjective data were collected using controller and observer questionnaires. Complete data were obtained for 22 measures and partial data were obtained for 5 measures. Summary measures were presented, in different levels of detail, for the overall simulation and by sector and specific time period. The data provided a meaningful representation of the radar operational position and a partial representation of the radar associate position.

Several limitations and constraints on the data were discussed. Refinements to baseline measures and changes in methodology were recommended. It was also suggested, for future baseline studies, that additional simulation runs be completed. Advice was provided on using baseline data to make comparisons between systems.

17. Key Words

En Route Air Traffic Control
Baseline Data
Plan View Display
Air Traffic Controller Performance Measurement

18. Distribution Statement

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	56	

Blank Pages

ACKNOWLEDGEMENT

The following individuals provided invaluable assistance in the preparation of this document:

Mike McAvoy, Chicago ARTCC
George Runner, Chicago ARTCC
James Karlovitch, Dallas/Fort Worth ARTCC (NATCA)
Mike Williams, Dallas/Fort Worth ARTCC
John Wojcik, Oakland ARTCC

TABLE OF CONTENTS

EXECUTIVE SUMMARY

A study was conducted to provide baseline measurements of today's en route system consisting of the Host Computer System, the Plan View Display, and the M1 console. Four operational constructs were specified as key to Air Traffic Control (ATC) operations: Safety, Capacity, Performance, and Workload. Through analysis, an additional operational construct, Usability, was derived. A sixth construct, Simulation Fidelity, was included to account for accuracy in the ATC simulations on which this report is based. Across these 6 operational constructs, 29 measurements were identified.

In order to obtain data on these measurements, an ATC simulation platform was used based on four sectors of Washington Air Route Traffic Control Center (ZDC) airspace. A scenario, predicated upon ZDC System Analysis Recording (SAR) data for a 90th percentile day for traffic volume, was used.

Objective data were reduced from SAR and Amecom tapes. Subjective data were collected using controller and expert observer questionnaires. Controller workload was measured using the Air Traffic Workload Input Tool, and keyboard inputs were reduced from SAR tapes. Complete data were obtained for 22 measures, and partial data for 5 measures with 2 remaining for further study.

Statistics were reported at several levels of granularity. A measurement summary provided system data aggregated across the four sectors. Some summary measures were split out by sector, some were reported only at the sector level, and others were reported as time-based data. Data from the study provided a meaningful representation of the radar controller operational position and a partial representation for the radar associate position.

Several limitations and constraints with the methods used to collect the currently available data on the baseline system were identified. Refinements to baseline measurements were recommended, and new measures associated with the radar associate controller and sector team operations were more fully defined. Changes to the methodology included the need to use the Target Generation Facility to expedite SAR data reduction with a future simulation. Additional simulation runs are also needed to stabilize the data and attenuate some of the variability attributable to control technique.

Plans for a future simulation and field activity to obtain baseline data on the complete set of measurements were described. These plans included training a new en route user team on the same ZDC airspace and local procedures.

Guidance on using the baseline measurements to verify the effectiveness and efficiency of a future system was presented. This included a process for merging quantitative statistics with controller expert opinion to determine comparability of the baseline and future systems. Statistical equivalency was defined on the basis of traditional descriptive and inferential parametric and non-parametric statistics. These data could be pertinent in mitigating risk associated with the acquisition of future systems.

1. INTRODUCTION.

As the FAA moves into the 21st century, new en route automation programs are being specified, prototyped, developed, tested, and deployed. These new systems will replace or augment systems currently in use in the field. In order to provide for the continued safe, orderly, and expeditious flow of air traffic, a suite of measurements have been developed. These measurements will define and quantify the level of operational efficiency and effectiveness of today's en route air traffic control (ATC) system.

This report identifies and defines a proposed suite of measurements and provides quantitative data on today's system. The key components comprising today's en route system are the Host Computer System (HCS), the Plan View Display (PVD), and the M1 console.

2. PURPOSE.

The purpose of this baselining effort was to identify measurements and, to the extent possible, collect data pertaining to ATC efficiency and effectiveness. These data were obtained by having en route controllers use the current HCS/PVD/M1 system to control realistic, simulated air traffic.

Air Traffic Requirements identified four high level operational constructs on which to base the efficiency and effectiveness of the ATC system: Safety, Capacity, Performance, and Workload. A fifth operational construct was derived in the course of this analysis: Usability. An additional construct was added to measure the representativeness of the ATC simulation: Simulation Fidelity. These constructs were defined as follows:

a. Safety represented the extent to which system-induced variables maintained, enhanced, or degraded relative safety (e.g., number of system errors, conflict alerts).

b. Capacity provided a measure of traffic through a specific section of airspace during a specified time period. Capacity changed as a function of controller, pilot, or system variables.

c. Performance involved controller interaction with the system through computer-human interfaces and included such data as number of data entries.

d. Workload represented cognitive and physical task requirements, along with actual time constraints, placed on performance. Additional measures were captured which provided a subjective index for individually perceived workload across time.

e. Usability consisted of performance envelopes associated with various aspects of the controller workstation, such as assessed using anthropometric models. Also included were user opinions regarding the acceptability of controls, displays, and other equipment items.

f. Simulation Fidelity represented characteristics of the air traffic mix, as well as the perceived fidelity of the simulation scenarios.

From each of these constructs, a set of baseline measurements was derived for which objective and subjective data could be obtained. Objective data were measurements that were pertinent to the ATC mission and realistic concerning ATC operations. Subjective data were obtained from controllers and observers and represented subject matter expert opinions and perceptions.

The measurements collected during the simulations provide indices of relative levels of operational acceptability and cannot be used in isolation. Variations in reported values must be analyzed in context with associated constructs to derive possible implications. Any other use of data from this analysis might prove misleading and invalid.

3. APPROACH.

Discussion of the approach is organized into a brief summary followed by a detailed description of the methodology.

3.1 SUMMARY.

The approach to PVD baselining was defined in terms of Full Performance Level (FPL) controllers working a high traffic volume through sectors of airspace determined to be representative of the National Airspace System (NAS). The Federal Aviation Administration (FAA) Technical Center Host PVD Display Computer Channel (DCC) Laboratory served as the platform that included use of the Host Dynamic Simulation (DYSIM) system for target generation.

The airspace selected was from the Washington Air Route Traffic Control Center (ARTCC), or ZDC. Two low altitude sectors (Sector 26, Sampson and Sector 27, Liberty) and two high altitude sectors (Sector 38, Tar River and Sector 35, Wilmington combined with Sector 09, Dixon) were used. Twelve controllers participated in this study. The sector radar controller (R) and radar associate or data controller (D) operational positions were staffed by current ZDC FPL controllers. Each sector was assigned two simulation pilot positions that were staffed by some of the ZDC controllers who had recent DYSIM pilot experience.

A scenario with two adjacent low altitude sectors (26 and 27) was used for familiarization with the baselining platform. The baseline simulation used two airspace configurations: adjacent low and high altitude sectors (26 and 38), and non-adjacent low and high altitude sectors (27 and 35).

The baseline simulation used a 90th percentile day for traffic volume, as defined in FAA Order 7210.46. This volume was sufficient for controllers to functionally exercise the HCS/PVD/M1 system. Special events and unscripted pilot requests were kept at a minimum.

During the actual simulation runs, measurements relating to system Safety, Capacity, Performance, Workload, Usability, and Simulation Fidelity were recorded. Objective data were automatically collected as System Analysis Recording (SAR) data, Air Traffic Management Program (AMP) data, and Amecom voice recordings. Manual objective tallies were recorded by

ATC specialists based upon their observations of ATC activities. Subjective data were collected by using expert observer logs, controller workload measurement rating tools, and other questionnaires. Videotaping was also utilized to record all controller activities.

At the end of testing, all data were reduced, compiled, and analyzed by ACD-350 and ACN-300. A Quick-Look Test Report was generated one week after test completion to provide an early look at preliminary test data and results. This report also presented a summary of the test. A Draft Final Test Report package was completed 45 days (March 21, 1995) after completion of the Quick-Look Report. The data in the current report represent an extrapolation from, and expansion to, the initial findings.

Two copies of all data were generated; one copy is being held by ACN-300 and the other has been given to ATR-320. All system tapes are being archived at the FAA Technical Center. Operational expertise in subsequent measurement definition, data collection, analysis, and interpretation was provided by an FPL controller from Dallas/Ft. Worth ARTCC and an area supervisor from Chicago ARTCC.

3.2 METHODOLOGY.

There were several major components of the baselining methodology used in this initial effort. These components consisted of the ZDC controllers and other ATC specialists who participated in this activity, the ZDC sectors, the simulation scenarios, the laboratory platform, the data collection schedule, the simulation runs, and the objective and subjective measurements taken.

3.2.1 Study Participants.

In association with the use of ZDC airspace, twelve ZDC controllers participated in this study. These controllers were current and knowledgeable of the airspace used in the simulation and staffed the sector R and D operational positions. Each sector was assigned two simulation pilot positions. Those ZDC controllers having recent DYSIM pilot experience staffed the pilot positions and controllers rotated through all positions. Two ATC specialists who were familiar with Operational Test and Evaluation procedures served as expert observers.

3.2.2 ZDC Airspace.

The airspace used in the simulation was from ZDC. Two low and two high altitude sectors were selected, as follows:

a. Sector 26, referred to as Sampson, is a low altitude sector from 11,000 feet (ft) to 23,000 ft that borders Jacksonville ARTCC. Sampson is completely underlaid with terminal airspace. Sampson interfaces with the following approach control facilities: Fayetteville, Raleigh, Seymour Johnson, Wilmington, and Patuxent River. A large part of the Sampson traffic comes from Raleigh airport southbound departures.

b. Sector 38, referred to as Tar River, is a high altitude sector of 24,000 ft and above which generally has northbound traffic. The Washington metropolitan airports (Dulles, National, and Baltimore/Washington) makes up most of the arrival flow for Tar River. This sector transitions Raleigh-Durham airport departures to the south and east from Rocky Mount and Sampson sectors to high altitude stratum. Other major traffic flows come southbound from New York metropolitan, New England, and Philadelphia airports.

c. Sector 27, referred to as Liberty, is a low altitude sector from 11,000 ft to 23,000 ft that borders Atlanta ARTCC. Liberty interfaces with Greensboro, Raleigh-Durham, and Fayetteville approach control facilities. This sector has numerous traffic flows, but a large amount of the flow comes from: Raleigh airport departures to the west and north; Raleigh airport arrivals from the southwest and south; Charlotte airport departures to the north and east; and Charlotte airport arrivals from the east. Liberty also works military traffic from Pope Air Force Base.

d. Sector 35, referred to as Wilmington, is combined with sector 09, referred to as Dixon. This is a high/ultra high altitude sector at 24,000 ft and above with a high volume of en route, north-south corridor traffic. Traffic flows are mostly from Miami, Orlando, Raleigh, JFK, and Philadelphia airports.

3.2.3 Simulation Scenarios.

The air traffic patterns and airspace characteristics were representative of the local adaptation of ZDC sectors. To ensure repeatability of baselining conditions, three ZDC DYSIMs were prepared. A scenario with adjacent sectors 27 (low) and 26 (low) was used for familiarization with the baselining platform. Adjacent sectors 26 (low) and 38 (high) and non-adjacent sectors 27 (low) and 35 (high) were used for the actual baseline simulation runs.

The baseline simulation runs used a 90th percentile day for traffic volume, which was determined to be sufficient for controllers to functionally exercise the HCS/PVD/M1 system. Special events and unscripted pilot requests were kept at a minimum in case such events would detract from the repeatability of the baselining scenarios.

The simulations were built from actual SAR flight data taken from ZDC (September 1992 time frame). The simulations were verified and rated by a current ZDC controller, then tested in the FAA Technical Center laboratories. There was a low, medium, and high intensity simulation for each sector combination build. The following simulations (sector numbers) were available for the PVD baseline testing:

a. 26, 27 low intensity adjacent
b. 26, 27 medium intensity adjacent
c. 26, 27 high intensity adjacent
d. 27, 35 low intensity non-adjacent
e. 27, 35 medium intensity non-adjacent
f. 27, 35 high intensity non-adjacent

g. 26, 38 low intensity adjacent
h. 26, 38 medium intensity adjacent
i. 26, 38 high intensity adjacent

Though all levels of intensity were available, only the high level intensity simulations were used, since the traffic ranged from low to high intensity during the course of each run. This gave controllers an adequate variation in task load for data collection purposes.

The PVD laboratory environment utilized for the simulation was realistic compared with a typical en route facility. The volume of traffic was sufficient to exercise the controllers' abilities, and the controllers were current and knowledgeable with the airspace used in the simulation. These factors contributed to a sound baseline and allowed the test team to collect credible data samples throughout the simulation period.

3.2.4 Laboratory Platform.

The PVD laboratory served as the platform that included use of the Host DYSIM system for target generation. The DYSIM provided effective control of the small-sized simulation and ensured repeatability of the scenarios.

Communications maps were built on the Amecom system. DYSIM pilots provided both air/ground and ground/ground (e.g., adjacent sectors, approach control facilities) communications to the controllers.

The Air Traffic Workload Input Technique (ATWIT) (Stein, 1985), used by controllers to rate workload, was incorporated into the PVD and the D Computer Readout Display (CRD) equipment. Every 4 minutes, an aural indication prompted the controller for an ATWIT. The R entered a workload rating from 1 (low) to 7 (high) on the PVD keyboard, and the D entered a rating on the CRD keyboard.

Small video cameras were strategically located to record sector activities. The videotapes were subsequently reviewed to augment these data analyses.

3.2.5 Data Collection Schedule.

Data collection began on January 10, 1995, with a pre-test briefing to the participating controllers and expert observers. The goal was to complete 3 simulation runs a day for 3 days. Due to FAA Technical Center PVD DCC Laboratory hardware problems, the first simulation run was aborted. After the problems from the first run were remedied, all subsequent simulation runs were successful. Testing was completed on the afternoon of Thursday, January 12, 1995.

3.2.6 Simulation Runs.

Familiarization simulations were executed 30 minutes before the start of the first 2 days of testing to help the controllers become acclimated.

There were a total of eight simulation runs. Three of the runs contained 1 hour of data and 5 of the runs captured 1 1/2 hours of data. The 1 1/2-hour run was a 2-sector simulation with ZDC non-adjacent sectors Liberty (27) and Wilmington (35). The 1-hour run was a 2-sector simulation with ZDC adjacent sectors Sampson (26) and Tar River (38). It was the first portion of a scenario originally planned for 1 1/2 hours that was shortened to accommodate scheduling changes. Controller assignments to sector and pilot positions were changed between runs.

3.2.7 Objective and Subjective Measurements.

Objective data were recorded on SAR and AMP tapes. NAS software modifications permitted DYSIM flights to be recorded on the AMP tape. Objective data were manually recorded, when necessary, through expert observer tallies. These measures focused on quantifying traffic volume, flight duration, and traffic characteristics in each sector. Another goal for recording objective data was to determine the input/output activity at each sector position to measure how each controller used the system. The objective data recorded captured all NAS activity to provide supplemental information that may assist in explaining possible discrepancies or anomalies.

Subjective data were gathered through pre-run and post-run questionnaires. Real time subjective data were recorded using ATWIT and specially scripted expert observer logs. Pre-run questionnaires recorded the experience levels, perceptions of air space and traffic characteristics, and user preferences of each participating controller from ZDC. Post-run questionnaires, real time ratings, and expert observer logs recorded perceived workload, capacity, controller performance, realism, traffic complexity, and system performance.

The following is a summary of the data types that were collected for this testing activity:

a. Pre-run controller questionnaires/background survey
b. Post-run controller questionnaires
c. Post-simulation controller survey/final questionnaire
d. Video tape (with audio)
e. Amecom audio tape from communications system
f. SAR tapes
g. AMP tapes
h. Real time controller workload ratings (ATWIT)
i. Real time expert observer logs
j. Post-run expert observer ratings of controller performance
k. Communication counts (expert observers)

The definitions for each of the baseline measures and their rationale for use in baselining today's PVD system are as follows:

a. Safety

1. Operational Error Rate was a basic safety measure representing loss of applicable separation minima. Operational errors were analyzed to determine the extent, if any, of system-induced causes.

2. Conflict Alert Rate was a system-initiated display derived from HCS tracking data, warning the controller of potentially imminent aircraft-to-aircraft conflicts. Possible sources of variation may have been due to the ease/difficulty for the controller to recognize potential future conflicts and the ability to maintain data blocks with current altitude assignments.

3. Use of Halo (J-Ring) was a controller-initiated display. It surrounded an aircraft target symbol with an adapted radius to aid in visual judgement of lateral and longitudinal separation. It was also useful as an emphasis tool and memory aid. Possible sources of variation may have been related to difficulty in using the display to judge aircraft separation or differences in controller ability to visualize spacing.

4. Data Block Offset/Leader Length was a controller-initiated function that oriented the data block by altering leader length and/or direction. It was used to maintain unimpeded readability of critical PVD data. Data Block Offset was an essential workload component, and its usage would have increased with traffic volume and associated data block overlap. Total usage was recorded in 12-minute segments, by sector. A secondary use was as a memory aid, that is, zero leader length to indicate transfer of communication. However, counts of controller message inputs excluded "slant zero" entries that were typically not associated with PVD data readability.

5. Other Safety-Critical Issues were expert observer comments on system safety issues and deficiencies. The observer logs were used to capture additional safety concerns not otherwise recorded.

b. Capacity

1. Aircraft Under Control was a basic capacity measure. At the measurement summary and sector summary levels, it represented a tally of traffic under track control. When treated as a time-based measure, it represented the total number of aircraft under track control by 12-minute segments.

2. Average Time in Sector was a measure of sector efficiency. Increased time in sector may have indicated less efficient movement of aircraft in the airspace.

3. Altitude Assignments Per Aircraft provided a ratio of total altitude assignments to number of aircraft under control. It was an indicator of the relative efficiency of aircraft movement through the sector. Controllers commonly relied on vertical separation in preference to vectoring solutions as perceived workload and complexity dictated. This resulted in level-offs

and climb/descent delays. A decrease in altitude assignments with a corresponding decrease in climb/descent delays would have indicated greater efficiency. An increase in altitude assignments with a corresponding increase in climb/descent delays and level-offs would have indicated less efficiency.

c. Performance

1. R Data Entries was a relative measure of data entry workload for the radar position. R Data Entries were counted per message type. Message types are specified in NAS-MD-311. The distribution of data entries could have shifted between the R and D positions. A qualitative analysis will be required to determine the source of workload variations.

2. R Data Entry Errors was a relative measure of data entry effectiveness. Significant variations have difficult message syntax, awkward entry device layout, or other possible factors.

3. D Data Entries was a relative measure of data entry workload for the radar associate or data position. D Data Entries were counted per message type. Message types are specified in NAS-MD-311. The distribution of data entries could have shifted between the R and D positions. A qualitative analysis will be required to determine the source of workload variations.

4. D Data Entry Errors was a relative measure of data entry effectiveness. Significant variations have difficult message syntax, awkward entry device layout, or other possible factors.

5. Timed Performance of Functions was the time required for controller input actions for Host computer messages. For this report, initial time data were taken from tentative workload data and will have to be confirmed.

6. Number of Altitude, Speed, and Heading Changes represented the efficiency of sector operations in terms of total number of clearances issued in these three categories and the proportion of each type. Significant variation would show that something had changed in the way controllers handle traffic. Counts were based upon aircraft-related data entries at the DYSIM pilot positions.

7. ATC Services was a measure of the quality of ATC services. It was used as an indicator of system usability. Measures were taken from the post-run controller questionnaire. The specific items comprising the measure were:

a) ATC services from the pilot's perspective.
b) Self-judgement of quality of ATC work.

8. Human Capabilities for ATC was a measure representing human capabilities used by the controller in performing ATC functions. Ratings were made by expert observers as part of their post-run ratings and served as indicators of operator efficiency/effectiveness. The specific items comprising the measure were:

a) Communicating/Informing
b) Managing Multiple Tasks
c) Technical Knowledge
d) Reacting to Stress
e) Maintaining Attention/Vigilance
f) Prioritizing
g) Maintaining Safe/Efficient Flow
h) Adaptability/Flexibility
i) Coordinating

d. Workload

1. Workload Per Aircraft was a measure that estimated the amount of workload expended per aircraft. It represented average subjective ATWIT responses versus number of aircraft tracked over 12-minute segments. Subjective workload ratings corresponded closely to number of aircraft tracked throughout the baseline scenarios. Workload Per Aircraft was measured separately for R and D.

2. Average Workload was the mean subjective workload reported by controllers, by sector, across the entire simulation. Workload was measured using the ATWIT. Workload was defined as a human response to the demands or task loads produced by the airspace system. Human response consisted of observable control actions and cognitive activity. Average Workload was measured separately for R and D.

3. Post-Run Workload was a measure that evaluated controller average workload for the scenario as part of the post-simulation run questionnaire. The rating scale ranged from 1 (low) to 8 (high). Post-Run Workload was measured separately for R and D.

4. Communication Actions was a measure that detected changes in communication workload needed to control aircraft. It provided a ratio of total sector communications versus number of aircraft tracked for 12-minute segments. Increased communications per aircraft may have indicated a less efficient automation interface. Conversely, increased communications per aircraft may have represented greater latitude on the part of controllers to maneuver aircraft and initiate actions.

5. Data Entry Workload was a measure that detected changes in average Data Entry Workload required to control aircraft. It compared the average number of data entries versus average number of aircraft tracked per 12-minute segment by sector.

6. Between Sector Coordination was a basic measure of sector coordination workload. Possible sources of variation could have included difficulty in completing sector tasks.

e. Usability

1. Strip Bay Management was an assessment of the ergonomics associated with strip bay management for the R and D operational positions. Five flight strip measures (one set for each of the R and D positions) were collected on: reaching, viewing, marking, inserting/ removing, and angling.

The score for each part of the evaluation was derived from either the number of flight strips that were accessible to reach, view, or mark or whether specific strips in different strip bay areas were subjectively rated as easy to insert, remove, or angle. Scores for each of the five components were summed to yield a total score for the HCS/PVD/M1 system (expressed as a percentage of the total possible score). These data were based on the performance of one controller. Review of the technique and additional data collection will be required to create an adequate baseline.

2. Within-Sector Coordination assessed the extent to which the existing system supports some aspects of teamwork between R and D. A static evaluation was designed that considered the number of flight strips both R and D could jointly access. Another component was the ability, given the ergonomics of the console, of the D to read and point to data blocks on the PVD from a normal seated position. The accessibility of information and devices for the handoff, or tracker, controller was also evaluated, as were any impediments to spoken communication between controllers. Scores on each component test were summed and expressed as a percentage of the total possible score. These data were based on the same performance of one controller. Review of the technique and additional data collection will be required to create an adequate baseline.

3. HCS/PVD/M1 System were measures of the usability of the system as rated by controllers. The specific items comprising these measures on the post-simulation controller questionnaire were as follows:

a) Flight Progress Strip Access
b) Flight Progress Strip Read/Mark
c) Ease of Access of Controls
d) Operation of Controls Intuitive
e) Keyboard Ease of Use
f) Radar and Map Displays Ease of Reading
g) Radar and Maps Displays Ease of Understanding
h) Workstation Space
i) Equipment, Displays, and Controls Support Efficient ATC
j) Equipment, Displays, and Controls Impose Limitations
k) Overall Effectiveness of Equipment, Displays, and Controls
l) Overall Quality of Interaction with Equipment

f. Simulation Fidelity

1. Traffic Characteristics was a measure representing the number of flights, type of flight (arrival, departure, and overflight), and type of aircraft (jet or propeller). It was used as a characterization of the simulation scenario.

2. Perceived Representativeness was a measure of the controllers' perceived fidelity of the simulation scenarios for the four sectors. It was used as a check on the realism of the simulation. The items comprising this measure on the post-run controller questionnaire were as follows:

a) Realism
b) Technical Problems
c) Problem Difficulty

4. MEASUREMENT SUMMARY DATA.

The summary of all measurements aggregated across all sectors and the corresponding simulation runs is shown in appendix A, table 1.

Measurement summary data provide system-level statistics for overall baseline HCS/PVD/M1 operations. The table includes all measurements used in this baselining effort. For some measurements, table 1 presents both the aggregated data and refers to more detailed sector information contained in tables 2 through 15, appendix A. This additional detailed sector information is intended to augment the aggregate data, that is, to assess trends between high and low altitude sectors. For certain measurements, table 1 indicates that aggregate data are not meaningful and refers to other tables containing the pertinent data.

For the measurement of Timed Performance of Functions, table 1 refers to table 5 for controller action times associated with NAS message entry categories.

Table 1 includes subjective questionnaire data collected from controllers and expert observers. Rating averages are reported separately for R and D operational positions. Subjective data collected at the end of the baseline simulation study are reported at only the measurement summary level, and not split out by sector. Other subjective data were collected after each run and are reported by sector in table 6.

For the measurements of Strip Bay Management and Within-Sector Coordination, table 1 presents the respective index score. These measurements are reported at only the summary level and are based upon a trial run using one controller.

5. SECTOR SUMMARY DATA.

Sector summary data, which are averages for each sector across simulation runs, are shown in appendix A, table 2. For some measurements, references are made to other tables providing additional or decomposed data. For example, the measurements for R and D are cross-referenced to table 4, in appendix A, for frequency of NAS message entries per sector.

An examination of data reported in table 2 should consider that measurements for sectors 26 and 38 are based upon 60-minute simulation scenarios, whereas sectors 27 and 35 used 90-minute scenarios. Given these time differences, and the qualitative differences between sectors, direct comparisons between sectors cannot be made.

6. TIME-PHASED SECTOR DATA.

For some measurements, it was operationally meaningful to compile statistics at the sector level on the basis of time. Twelve-minute intervals were used to aggregate time-phased data. For example, ATWIT data were collected at every 4-minute rating input period, and averaged within each 12-minute period. These data are shown in appendix A, tables 8 through 15 for the four sectors. The tables indicate the number of time segments used in the statistics. For most sectors, no communications data were reduced for the last time segment.

7. CONCLUSIONS.

The conclusions based upon the results and findings of this study discuss methodological considerations with the present study. Guidance on comparisons of Plan View Display (PVD) baseline data with future en route systems and plans for further PVD baselining simulation runs are reviewed in appendixes C and D.

7.1 METHODS.

Methodological considerations associated with the present study include the limitations and constraints of the current simulation, refinements to the baselining methodology, and definition of further baselining requirements.

7.1.1 Limitations and Constraints of Current Simulation.

During the post-hoc analysis of the data and review of the draft PVD Baseline Research Report (dated March 21, 1995), some limitations and constraints were identified.

In the conduct of the simulation runs, for methodological reasons, there was an unequal number of runs made across sectors. Sectors 27 and 35 had five runs each, while sectors 26 and 38 had three runs each. Comparison of various baselining measures showed that statistics trended toward increased stability as the number of runs increased. The limited and unequal number of runs across sectors resulted in an unbalanced experimental design, and somewhat less stable data for two sectors. The existing operational data provide confidence intervals on which some statistical inferences may be based.

The duration of the simulation scenario runs varied across sectors. Sectors 27 and 35 had a nominal run duration time of about 90 minutes, and some runs were for 80 minutes. Sectors 26 and 38 had nominal run duration time of about 60 minutes. Sectors 26 and 38 had 1 run of 94 minutes. Different durations of simulation scenario run times across sectors necessitated additional calibrations to compare data and identify trends between sectors.

The use of the 90 percent traffic volume scenario placed limited stress on system operation. For example, it did not necessitate use of a third controller in a handoff/tracker operational position. This volume is representative of actual Washington Air Route Traffic Control Center (ARTCC), or ZDC, air traffic control (ATC) operations. The scenarios deliberately did not include or induce other potential anomalies associated with air traffic operations. It is recognized that other levels of system capacity should be tested that include traffic volumes associated with use of a third controller. It may also be important to test a single controller working combined radar controller (R) and the radar associate or data controller (D) operational positions under low traffic volumes.

The provision for use of video recording was intended to provide a historical record of each of the simulation runs to support later data analysis. Review of the video tapes demonstrated that closer positioning of cameras to controllers and workstations would be needed to provide sufficient clarity to avoid missing or misinterpreting controller actions and communications.

Two ATC specialists served as observers and were asked to make general observations of the controllers during each of the simulation runs and manually record their comments on log sheets. Review of observer data showed that their comments consisted of qualitative notes on controllers' actions and control techniques. It was difficult in the post-hoc analysis to correlate observer data with video recordings; recorded comments were not always time-stamped.

The analysis of data entry errors by the R and D operational positions, as reported in sections 4 and 5, were based on tallies across the simulation scenarios. Current Data Analysis and Reduction Tool (DART) and Air Traffic Management Program (AMP) data reduction capabilities do not accommodate the decomposition of errors by message types, as specified in National Airspace System (NAS)-MD-311 (FAA, 1991). This precluded examining differences in error rates across different message formats.

The tables described in sections 4, 5, and 6, and shown in appendix A represent a comprehensive array of baseline data for the R position and a partial representation of the D position. The methodology and the manner of data collected for the D position and for sector team operations have also been more fully defined.

7.1.2 Refinements to the Baselining Methodology.

Through this post-hoc analysis, several refinements were identified to further define the baselining methodology. Some of these refinements were provided by Air Traffic, and others were identified in the course of analyzing the baseline data.

The original suite of measurements contained in the draft report of March 21, 1995 was re-engineered in part to support Air Traffic's determination of four primary operational constructs.

These constructs were Safety, Capacity, Performance, and Workload. An additional operational construct, Usability, was added through data analysis. Another construct, Simulation Fidelity, was added to assess the representativeness of the ATC simulation. Several measurements were identified for each of the operational constructs.

An important baseline measurement for Capacity is Aircraft Fuel Consumption. This is an indicator of sector efficiency and could be based upon sector boundary crossing time in contrast to track control time. Fuel consumption could be measured according to average pounds of fuel consumed for all aircraft, by sector. Models would need the capability to import System Analysis Recording (SAR) data.

Added data are needed for the Performance measure of Timed Performance of Functions. Measurements need to be based upon a larger number of controllers.

Additional subjective measurements for Usability have been identified. These are the acceptability of display coding, ease of trackball use, and the quality of the communications interface. These items are other indicators of the effectiveness of the user interface.

An ergonomic measurement for Usability is reach envelope. This represents accessibility of controls and flight strips. Data have been reported elsewhere on the M1 flight strip reach envelope.[1] In addition, environmental factors associated with usability are noise, lighting, and electromagnetic emissions. These factors provide objective measurements on the quality of the work environment and should be considered for future baseline efforts. New ergonomic measures were developed for the R and D operational positions (i.e., Strip Bay Management) and sector team operations (i.e., Within-Sector Coordination). These methods are described in section 3 and will need further refinement.

With respect to the 6 operational constructs, a total of 29 measurements were identified. This report provides complete data for 22 measures, and partial data for 5 measures, with 2 remaining for further study.

In future PVD baseline simulations, scenario duration times should be 90 minutes for all sectors. This will maintain the timing of traffic surges associated with the 90 percent traffic volume. It will also mitigate a need to operationally and statistically calibrate data between sectors having different run duration times. Data contained in this report may be merged with future simulations, depending upon statistical comparisons and consideration of different study participants.

Incorporation of command entry keystroke data can be accommodated as static data augmenting R and D data entries. Keystroke data could be used to weight message inputs and could be derived from NAS-MD-311.

[1] A paper entitled "Comparative Analysis of Flight Strip Reach Envelopes: PVD and DSR" was prepared by J. Galushka and R. Mogford for ATR-320 in March, 1995 and is available from ACT-530. The approach employed in this paper should be incorporated into the Strip Bay Management assessment procedure in future PVD baseline study work.

The responsibilities, instructions, and logs used by expert observers need to be more specific. Additional structure in defining the types of general and time-based quantitative and qualitative data to be manually recorded by these observers will ensure the logging of useful information. Post-run debriefing of observers and controllers participating in each simulation run will increase the fidelity of the data collected through automated means, such as by identifying dynamic simulation anomalies and explaining ATC events and control techniques.

In order to fully establish a statistically and operationally valid baseline, there is a need for a minimum of 10 simulation runs per sector. The data contained in this report demonstrate increased stability in the measures afforded by using five simulation runs compared to three runs. However, even with five runs, examination of the measurement data shows some variability and skewness attributable to such factors as control technique. In comparison, the United Kingdom's New En Route Control (NERC) ATC studies have indicated a requirement for 17 simulation runs to achieve stability based upon their data analysis (Goillau, Kelly, Finch, & Arnold, 1994). In the data contained in the NERC studies, it is possible that intervening variables drive their requirement for a greater number of simulation runs to achieve stability in the data. These intervening variables potentially include the small number of study participants and the limited operational background of these individuals working as Full Performance Level (FPL) controllers.

<u>7.1.3 Definition of Further Baselining Requirements</u>.

Future PVD baseline simulations should use the Target Generation Facility (TGF). A direct benefit will be speedier data reduction attributable to TGF automated tools. The TGF will also support pilot transfer of aircraft control and air/ground communications as aircraft traverse adjacent sectors.

During the preparation of this report, the need for additional automated tools was identified to expedite data reduction and analysis. These tools would be used off-line commencing after completion of the first simulation runs and in parallel during the course of the remaining simulation runs. In this manner, data could be presented in a timely and precise manner shortly after conclusion of the last simulation run. Additional DART programs are needed to efficiently extract all pertinent data from SAR and AMP recordings. A program to quickly extract pertinent communications data from Amecom tapes is needed. Excel spreadsheet(s) should be specified and formulated to accommodate and automatically analyze all measurements to produce the required statistics.

Part of the DART and AMP reductions of SAR data should include several refinements. Data entry errors by the R and D operational positions should be classified by message type, in accordance with NAS-MD-311. This additional information needs to be reduced from the SAR data to account for error rates and may identify certain error-prone entries.

Additionally, the methodology should capture aborted message entries. Controller use of the keyboard "Clear" and other keys would provide further quantitative data on message entries and errors. Changes to the methodology could include some automated recording of individual key presses, or zoomed-in video recording of the CRD and/or the keyboard.

References

Goillau, P. J, Kelly, C. J., Finch, W. F., & Arnold, M. (1994). CAER new en route control (NERC) reference system (NRS) interim trial report (Report No. DRA/CSS4/CTR/RPT/CD189/1.0). Farnborough, UK: CAA.

Stein, E. S. (1985). Air traffic controller workload: An examination of a workload probe (Report No. DOT/FAA/CT-TN84/24). Atlantic City, NJ: FAA Technical Center.

Federal Aviation Administration. (1991). Computer program functional specification, message entry and checking (PAMRI Model A4e1.2) (Report No. NAS-MD-311). Washington, DC: Federal Aviation Administration.

APPENDIX A
MEASUREMENT SUMMARY AND SECTOR DATA

TABLE 1. MEASUREMENT SUMMARY TABLE FOR ALL CONSTRUCTS

Construct	Variable	Description	Rationale	HCS/PVD/M1 Value[1]	Comment
Safety	Operational Errors	Loss of applicable separation minima (per FAA Order 7210.3K).	Basic safety measure.	Total number: 0	N/A
	Conflict Alerts	Host conflict prediction algorithm.	Warning of potential conflict.	Total number/run for all sectors:[2,3] 12	See table 2 for sector information.
	Use of Halo (J Ring)	Visual range display used to indicate separation.	Could indicate difficulties judging aircraft separation from display.	Total number/run for all sectors: 14.3	
	Data Block Offset and Leader Length	Separates data blocks for readability.	Access to critical information.	Total number of direction and length changes/run for all sectors: 217.8	
	Other Safety-Critical Issues	Observations of system safety deficiencies.	Capture additional safety concerns not otherwise recorded.	No controller or observer system safety concerns.	N/A
Capacity	Aircraft Under Control	Total number of aircraft under track control.	Basic capacity measure.	Total number of aircraft handled/run for all sectors: 250.8	See table 2 for sector information.
	Average Time in Sector	Time from handoff to handoff.	Basic capacity/ efficiency measure.	N/A[4]	

[1] All data reported are for the full run time of each traffic scenario which was one hour for sectors 26 and 38 and about 90 minutes for sectors 27 and 35.
[2] The term "for all sectors" indicates that the number reported was a sum of the results at the sector level (table 2).
[3] A score of .5 was given for each aircraft showing a conflict alert.
[4] "Not Applicable" indicates that it was not appropriate to report an average or sum across sectors for this variable.

TABLE 1. MEASUREMENT SUMMARY TABLE FOR ALL CONSTRUCTS

Construct	Variable	Description	Rationale	PVD/M1 Value	Comment
Capacity (Continued)	Altitude Assignment Efficiency	Ratio of total altitude changes and number of aircraft.	Detects efficiency in moving flights through airspace.	N/A	See table 2 for sector information.
Performance	R Data Entries	Total data entries and breakdown by category.	Measures effort required to make data entries into system.	Total entries/run for all sectors: 972.7	See table 2 for sector information and table 4 for breakdown by categories.
	R Data Entry Errors	Total data entry errors.	Detects data entry problems.	Total errors/run for all sectors: 47.6	See table 2 for sector information.
	D Data Entries	Total data entries and breakdown by category.	Measures effort required to make data entries into system.	Total entries/run for all sectors: 319.9.	See table 2 for sector information and table 4 for breakdown by categories.
	D Data Entry Errors	Total data entry errors.	Detects data entry problems.	Total errors/run for all sectors: 54.6	See table 2 for sector information.
	Timed Performance of Functions	Test time to complete pre-defined set of system functions in static test.	Evaluates efficiency of user interface.	N/A	See table 5 for breakdown of time by data entry type.
	Number of altitude, speed, and heading changes	Count of DYSIM pilot entries to control aircraft (in response to controller instructions).	Indicates user interface effectiveness.	Total number/run for all sectors: 371.8	See table 2 for sector information.

TABLE 1. MEASUREMENT SUMMARY TABLE FOR ALL CONSTRUCTS

Construct	Variable	Description	Rationale	PVD/M1 Value	Comment
Performance (Controller Questionnaire)	1. ATC Services (Pilot)	Measures of quality of service.	Indicates system usability.	Average rating (R/D[5]): 6.9/7.1	See table 6 for sector information.
	2. How well did you control?			Average rating (R/D): 7.5/6.9	
(Expert Observer Questionnaire)	1. Communicate/ Inform	Measures of controller performance as evaluated by expert observers.	Indicates system efficiency/ effectiveness.	Average rating (R/D): 6.8/6.3	See table 6 for sector information.
	2. Manage Multiple Tasks	Measures of controller performance as evaluated by expert observers.	Indicates system efficiency/ effectiveness.	Average rating (R/D): 6.9/6.2	
	3. Technical Knowledge			Average rating (R/D): 6.9/6.6	
	4. React to Stress			Average rating (R/D): 5.2/5.1	
	5. Maintain Attention/Vigilance			Average rating (R/D): 7.2/6.7	
	6. Prioritizing			Average rating (R/D): 7.0/6.4	
	7. Maintain Safe/Efficient Flow			Average rating (R/D): 6.9/6.7	
	8. Adaptability/ Flexibility			Average rating (R/D): 6.8/5.5	
	9. Coordinating			Average rating (R/D): 7.0/6.9	

[5] "R/D" indicates that the results shown are separate average ratings for the radar and data controllers to a post-run questionnaire. The ratings ranged from 1 (strongly disagree) to 8 (strongly agree).

TABLE 1. MEASUREMENT SUMMARY TABLE FOR ALL CONSTRUCTS

Construct	Variable	Description	Rationale	HCS/PVD/M1 Value	Comment
Workload	Workload per Aircraft	Ratio of subjective workload (ATWIT) and number of aircraft tracked.	Detects changes in subjective workload to control aircraft.	N/A	See table 2 for sector information.
	Average Workload	Average ATWIT workload per run.	Detects changes in subjective workload to control aircraft.		
	Post-Run Workload	Subjective workload as measured by questionnaire at the end of each run.	Detects changes in subjective workload to control aircraft.		
	Communication Workload	Ratio of total communications and number of aircraft.	Detects changes in communications needed to control aircraft.		
	Data Entry Workload	Ratio of total data entries and number of aircraft.	Detects changes in data entries needed to control aircraft.		
	Between-Sector Coordination	Number of communications needed for coordination.	Measure of communication workload for coordination.	Total count/run for all sectors: 39.3	
Usability	Strip Bay Management	Evaluation of strip bay activities (marking, posting, sorting, cocking, etc.).	Indicator of effectiveness of strip bay layout , strip format, and strip use.	Strip Bay Index: R - 54% D - 87%	See section 3.2.7 for discussion.

TABLE 1. MEASUREMENT SUMMARY TABLE FOR ALL CONSTRUCTS

Construct	Variable	Description	Rationale	PVD/M1 Value	Comment
Usability (continued)	Within-Sector Coordination (R & D Teamwork)	Indicator of how well system supports coordination.	Assessment of usability of system with regard to coordination.	Coordination Index: 60%	See section 3.2.8 for discussion.
	1. Flight Progress Strip Access	System usability measures.	Indicators of efficiency/ effectiveness of user interface	Average rating (all[6]): 5.4 (1.78)[7]	N/A
	2. Flight Progress Strip Read/Mark			Average rating (all): 6.0 (1.04)	
	3. Ease of Access of Controls			Average rating (all): 6.3 (1.07)	
	4. Operation of Controls Intuitive			Average rating (all): 6.3 (.75)	
	5. Keyboard Ease of Use			Average rating (all): 6.3 (.65)	
	6. Radar and Map Displays Ease of Reading			Average rating (all): 5.5 (1.68)	
	7. Radar and Maps Displays Ease of Understanding			Average rating (all): 5.7 (1.30)	
	8. Workstation Space			Average rating (all): 6.1 (1.00)	
	9. Equipment, Displays, and Controls Support Efficient ATC			Average rating (all): 5.6 (.90)	
	10. Equipment, Displays, and Controls Impose Limitations			Average rating (all): 5.1 (.67)	

6 "All" indicates that this rating was the average for all controllers in the Baseline Study as measured in a post-experiment questionnaire.
7 The numbers in parentheses are standard deviations.

TABLE 1. MEASUREMENT SUMMARY TABLE FOR ALL CONSTRUCTS

Construct	Variable	Description	Rationale	PVD/M1 Value	Comment
Usability (continued)	11. Equipment, Displays, and Controls Overall Effectiveness	System usability measures.	Indicators of efficiency/ effectiveness of user interface	Average rating (all): 5.3 (.78)	N/A
	12. Overall Quality of Interaction with Equipment			Average rating (all over 6 scales): 4.7 (1.45)	
Simulation Fidelity	Traffic Characteristics	Number of flights, type of flight (arrival, departure, en route), and aircraft type.	Characterization of simulation.	N/A	See table 2 for sector data.
	1. Realism	Perceived fidelity of simulation scenarios	Check on realism of simulation.	Average rating (R/D): 4.9/5.2	See table 2 for sector data.
	2. Technical Problems			Average rating (R/D): 3.9/2.8	
	3. Problem Difficulty			Average rating (R/D): 4.7/4.1	

TABLE 2. QUANTITATIVE SECTOR DATA

Construct	Variable	26	27	35	38	Comment
Safety	Operational Errors	0.0	0.0	0.0	0.0	N/A
	Conflict Alerts	3.3	5.7	0.8	2.2	See section 3.2.7 for discussion.
	Use of J Ring	5.0	5.4	1.2	2.7	N/A
	Data Block Offset and Leader Length	42.7	57.8	85.0	32.3	See tables 8-11 for time interval data.
	Other Safety-Critical Issues	0.0	0.0	0.0	0.0	See section 3.2.7 for discussion.
Capacity	Aircraft Under Control	57.3	63.2	81.0	49.3	See tables 8-11 for time interval data.
	Average Time in Sector (min)	11.1	9.8	13.8	7.8	See section 3.2.7 for discussion.
	Altitude Assignments per Aircraft	0.8	0.8	0.1	0.5	See tables 8-11 for time interval data.
Performance	R Data Entries	210.0	271.2	291.8	199.7	See tables 8-11 for time interval data and table 4 for category breakdown.
	R Data Entry Errors	11.0	12.6	12.0	12.0	See tables 8-11 for time interval data.
	D Data Entries	78.0	96.2	99.0	46.7	See tables 8-11 for time interval data and table 4 for category breakdown.
	D Data Entry Errors	11.0	24.4	10.2	9.0	See tables 8-11 for time interval data.
	Number of altitude, speed, and heading changes.	143.3	166.8	41.4	20.3	N/A

Note: All values, unless otherwise noted, are frequencies averaged across runs.

TABLE 2. QUANTITATIVE SECTOR DATA
(Continued)

Construct	Variable	26	27	35	38	Comment
Workload	R-Workload per Aircraft	0.3	0.2	0.2	0.2	See tables 8-11 for time interval data.
	D Workload per Aircraft	0.2	0.2	0.2	0.2	See tables 8-11 for time interval data.
	R Average Workload	3.8	3.2	3.1	1.8	See tables 8-11 for time interval data.
	D Average Workload	2.8	2.4	2.8	2.5	See tables 8-11 for time interval data.
	R Post-Run Workload	4.0	5.0	5.6	4.0	N/A
	D Post-Run Workload	3.3	5.2	4.4	3.0	N/A
	R Communications per Aircraft	2.8	2.5	2.0	2.5	See tables 8-11 for time interval data.
	Data Entries per Aircraft (R+D)	2.9	3.3	2.4	3.3	See tables 8-11 for time interval data.
	Between-Sector Coordination	8.3	13.6	7.4	10.0	N/A
Simulation Fidelity	Number of Arrivals	5.0	12.0	0.0	0.0	N/A
	Number of Departures	21.0	11.0	0.0	0.0	N/A
	Number of En Route	31.3	40.2	81.0	49.3	N/A
	Number of Jets	29.2	30.3	68.9	41.9	N/A
	Number of Propeller	28.1	32.9	12.2	7.4	N/A
	Scenario Length (min)	60	90	90	60	N/A

Note: All values, unless otherwise noted, are averaged across runs.

TABLE 3. QUANTITATIVE SECTOR STANDARD DEVIATION DATA

Construct	Variable	26	27	35	38
Safety	Operational Errors	N/A	N/A	N/A	N/A
	Conflict Alerts	1.04	2.77	0.84	1.26
	Use of J Ring	4.58	4.67	1.30	1.53
	Data Block Offset and Leader Length	20.79	25.90	16.32	3.51
	Other Safety-Critical Issues	N/A	N/A	N/A	N/A
Capacity	Aircraft Under Control	0.58	0.45	0.00	0.58
	Average Time in Sector (min)	1.77	0.65	1.24	1.49
	Altitude Assignments per Aircraft	0.26	0.08	0.02	0.07
Performance	R Data Entries	50.41	21.67	18.57	4.04
	R Data Entry Errors	3.00	6.19	5.96	2.65
	D Data Entries	23.07	7.95	1.22	8.50
	D Data Entry Errors	1.00	11.72	10.31	8.54
	Number of altitude, speed, and heading changes.	18.36	5.00	13.61	2.12
Workload	R Workload per Aircraft	0.08	0.07	0.01	0.05
	D Workload per Aircraft	0.09	0.06	0.09	0.19
	R Average Workload	1.89	0.62	0.95	0.72
	D Average Workload	1.41	0.42	0.70	0.26
	R Post-Run Workload	2.65	1.00	1.52	1.73
	D Post-Run Workload	2.52	1.10	1.52	1.73
	R Communications per Aircraft	0.38	0.08	0.14	0.36
	Data Entries per Aircraft (R+D)	0.85	0.36	0.23	0.04
	Between-Sector Coordination	7.50	7.20	4.30	4.40
Simulation Fidelity	Number of Arrivals	N/A	N/A	N/A	N/A
	Number of Departures	N/A	N/A	N/A	N/A
	Number of En Route	N/A	N/A	N/A	N/A
	Number of Jets	N/A	N/A	N/A	N/A
	Number of Propeller	N/A	N/A	N/A	N/A

TABLE 4. NAS MESSAGE ENTRIES PER SECTOR
FOR DATA AND RADAR POSITIONS

	Sector	26	26	27	27	35	35	38	38
	Controller	R	D	R	D	R	D	R	D
Message Type	Command								
Amend/delete FP data		0	0	0	0.6	0	0	0	0
Amend route/remarks	AM	0	0.7	0	0.4	0	0.6	0	1
Altimeter request	AR	0	0	0	0	0	0	0	0
Altimeter set	AS	0	0	0	0	0	0	0	0
EMSAW alert processing	CO	1.7	0	3	0	0.2	0	3.3	0
Departure message processing	DM	0	0.7	0	0	0	0	0	0
Beacon Code request	DQ	0	0	0	7.6	0	0	0	0
Flight Plan (FP) processing	FP	0	13	0	31.4	0	39	0	11.3
FP readout request	FR	0	3.7	0	0.2	0	0	0	1.7
General Information (GI) processing	GI	0	0	0	0	0	0	0	0
Hold message	HM	0	0	0	0	0	0	0	0
Range/Bearing readout	LA	2.3	0	4.6	0	3.8	0	1	0
Range/Bearing/Fix readout	LB	2	0	1.4	0	1.2	0	3.7	0
Fix/Time readout	LC	0	0	0	0	0	0	0	0
Route Fix/Time readout	LD	0	0	0	0	0	0	0	0
Route Fix/Time/Speed readout	LE	0	0	0	0	0	0	0	0
Cancel Mission Plan (MP)	MP	0	0	0	0	0	0	0	0
Progress report	PR	0	0	0	0	0	0	0	0
Automatic handoff processing	QA	0	0	0	0	0	0	0	0
Aircraft Beacon code processing	QB	0.3	0.7	0	0	0.6	0	1	0
CRD altimeters and PVD altitude limits	QD	0.3	0	0.2	0	0	0	0	0
FP readout request	QF	4.3	0	3.6	0	1.6	0	5.3	0
Hold message	QH	0	0	0	0	0	0	0	0
Initiate handoff	QN	108	0	208.8	0.2	286.6	0	165.3	0.3

TABLE 4. NAS MESSAGE ENTRIES PER SECTOR
FOR DATA AND RADAR POSITIONS
(Continued)

	Sector	26	26	27	27	35	35	38	38
	Controller	R	D	R	D	R	D	R	D
Message Type	Command								
FDB Request/supress	QP	10	1.3	10.6	3.2	2.4	0.4	7.7	1.7
Assign interim altitude	QQ	47.3	2.3	46.8	8.2	3.2	0	22.3	1
Report assigned altitude	QR	0	0	0.2	0	1	0.2	0	0.3
Amend assigned altitude	QT	0	0	1.4	2	0.6	0	0	0
Route display	QU	8.3	4.7	6.4	1.4	2.2	0	2	0.3
Drop track and remove strip	QX	0	0.3	0.6	1.4	0	0	0	0
Initiate handoff	QZ	3.7	14.7	2	8	5.8	0.2	5	5.3
Response to incorrect input action	REJECT	23	13.3	32	16	31.2	10	26.3	8.7
Transfer FP to ARTS	RF	0	0	0	0	0	0	0	0
Remove strip	RS	0	0.3	0	0	0	0	0	0
Cancel FP HOST only	RX	0	0	0	0	0	0	0	0
STEREO FP, abbreviated	SP	0	0	0	10.4	0	0	0	0
Strip request	SR	0	38.3	0	39.8	0	58.6	0	25
Terminate Beacon Code	TB	0	0	0	0	0	0	0	0
TEST device	TD	0	0	0	0	0	0	0	0
Upper winds request	UR	0	0	0	4.4	0	0	0	0
Enter weather data	WX	0	0	0	0	0	0	0	0

TABLE 5. MESSAGE ENTRY TIMES

HOST MSG TYPE	MESSAGE ACTION	INPUT TIME (SEC.)
AM	Amend/delete FP data	8
AM	Amend route/remarks	16
AM	Route readout	4
AR	Altimeter request	6
AS	Altimeter set	8
CO	EMSAW alert processing	5
DM	Departure message processing	5
DQ	Beacon Code request	5
FP	Flight Plan (FP) processing	40
FR	FP readout request	4
GI	General Information (GI) processing	5
HM	Hold message	8
LA	Range/Bearing readout	4
LB	Range/Bearing/Fix readout	4
LC	Fix/Time readout	4
LD	Route Fix/Time readout	4
LE	Route Fix/Time/Speed readout	4
MP	Cancel Mission Plan (MP)	6
MP	MP FP processing	40
PR	Progress report	8
QA	Automatic handoff processing	4
QB	Aircraft Beacon code processing	5
QB	PVD Beacon code processing	7
QD	CRD altimeters and PVD altitude limits	7
QF	FP readout request	4
QH	Hold message	8
QN	Initiate handoff	3
QN	/OK Accept handoff	3
QN	Accept handoff	2
QN	/OK Accept handoff	4
QN	Retract handoff	2
QN	FDB-position & distance	4

TABLE 5. MESSAGE ENTRY TIMES
(Continued)

HOST MSG TYPE	MESSAGE ACTION	INPUT TIME (SEC.)
QN	FDB-reposition	3
QN	FDB Request/supress	2
QN	Weather Report (ZAU)	2
QP	FDB Request/supress	4
QP	Delete A/C from metering list	6
QP	Point out-Initiating Sector	5
QP	Reposition List	2
QP	Distance Ref. Indicator	2
QQ	Assign interim altitude	5
QQ	Remove interim altitude	4
QR	Report assigned altitude	5
QT	Amend assigned altitude	5
QT	Start Coast Track	5
QT	Start Track	5
QU	Route display	4
QU	Track reroute	9
QX	Drop track and remove strip	5
QX	Drop track only	4
QZ	Initiate handoff	3
QZ	/OK Initiate handoff	3
QZ	Accept handoff	2
QZ	/OK Accept handoff	4
QZ	Retract handoff	2
QZ	Assigned altitude	5
QZ	FDB distance change	4
QZ	FDB position & distance change	4
QZ	FDB reposition	3
QZ	FDB request/suppress	2
RF	Transfer FP to ARTS	5
RS	Remove strip	4
RX	Cancel FP HOST only	4
SG	Conflict Alert processing	6

TABLE 5. MESSAGE ENTRY TIMES
(Continued)

HOST MSG TYPE	MESSAGE ACTION	INPUT TIME (SEC.)
SP	STEREO FP, abbreviated	13
SP	STEREO FP, processing	29
SR	Strip request	7
TB	Terminate Beacon Code	4
TD	TEST device	3
UR	Upper winds request	4
WR	Request/display weather data	2
WX	Enter weather data	16

TABLE 6. QUESTIONNAIRE DATA

Performance	Questionnaire Item	Sector 26 R	Sector 26 D	Sector 27 R	Sector 27 D	Sector 35 R	Sector 35 D	Sector 38 R	Sector 38 D
(Controller)	1. ATC Services (Pilot)	7.0	6.0	6.4	7.6	7.0	7.6	7.3	7.3
	2. How well did you control?	7.3	6.3	7.6	6.8	7.2	7.0	7.7	7.3
(Expert Observer)	1. Communicate/Inform	6.3	5.7	6.6	6.4	7.2	5.6	7.0	7.3
	2. Manage Multiple Tasks	6.0	5.7	7.8	6.0	6.2	5.8	7.7	7.3
	3. Technical Knowledge	6.3	6.0	7.0	5.8	6.4	7.4	7.7	7.0
	4. React to Stress	6.3	6.0	6.8	6.6	5.4	3.4	2.3	4.3
	5. Maintain Attention/Vigilance	6.3	5.7	7.6	6.4	7.4	7.0	7.3	7.7
	6. Prioritizing	6.0	5.7	6.6	6.6	7.8	5.6	7.7	7.7
	7. Maintain Safe/Efficient Flow	6.0	5.7	7.2	6.2	7.2	7.2	7.0	7.7
	8. Adaptability/Flexibility	6.0	5.7	6.8	6.0	7.2	5.4	7.0	5.0
	9. Coordinating	6.3	7.0	7.2	6.8	7.4	5.6	7.0	8.0
Simulation Fidelity (Controller)	1. Realism	4.3	3.7	5.0	5.6	6.4	6.6	3.7	5.0
	2. Technical Problems	5.0	3.7	3.6	2.2	2.4	2.6	4.7	2.7
	3. Problem Difficulty	4.0	3.7	4.4	5.2	6.0	4.6	4.3	3.0

TABLE 7. QUESTIONNAIRE STANDARD DEVIATION DATA

		Sector 26	Sector 26	Sector 27	Sector 27	Sector 35	Sector 35	Sector 38	Sector 38
Performance	Questionnaire Item	R	D	R	D	R	D	R	D
(Controller)	1. ATC Services (Pilot)	1.73	1.73	2.19	0.89	1.00	0.55	1.15	1.15
	2. How well did you control?	1.15	1.53	0.55	1.30	1.10	1.73	0.58	1.15
(Expert Observer)	1. Communicate/Inform	1.53	1.53	1.82	1.52	2.17	3.51	1.73	0.58
	2. Manage Multiple Tasks	1.00	1.53	1.30	1.00	1.30	3.42	0.58	0.58
	3. Technical Knowledge	1.15	1.00	0.71	1.30	1.34	1.34	0.58	1.00
	4. React to Stress	1.53	1.00	1.79	1.67	3.51	3.44	4.04	3.79
	5. Maintain Attention/Vigilance	1.15	1.53	1.52	1.67	1.82	1.58	3.06	0.58
	6. Prioritizing	1.00	1.53	2.07	1.52	2.28	3.36	2.52	0.58
	7. Maintain Safe/Efficient Flow	1.00	1.53	1.92	1.48	1.92	1.64	2.65	0.58
	8. Adaptability/Flexibility	1.00	0.58	1.92	1.58	1.92	3.29	1.73	4.36
	9. Coordinating	1.15	2.65	1.92	1.92	1.95	3.51	1.73	0.00
Simulation	1. Realism	3.06	2.52	1.22	1.67	0.55	0.55	1.15	1.00
Fidelity	2. Technical Problems	3.00	3.06	1.52	1.79	1.52	1.52	2.08	1.15
(Controller)	3. Problem Difficulty	2.65	2.52	1.14	1.48	1.00	1.34	2.08	2.65

TABLE 8. SECTOR 26 - 12-MINUTE INTERVAL DATA

Construct	Variable	1	2	3	4	5	6
Safety	Data Block Offset and Leader Length	2.7	7.7	15.0	6.7	10.7	0.0
Capacity	Aircraft Under Control	8.7	16.3	15.3	13.7	14.0	7.0
	Altitude Assignments per Aircraft	0.1	0.7	0.7	0.7	0.7	0.9
Performance	R-Data Entries	17.3	42.0	54.0	42.7	46.7	22.0
	R Data Entry Errors	2.7	2.7	2.0	1.0	2.3	1.0
	D Data Entries	13.0	10.3	16.0	15.3	22.0	4.0
	D Data Entry Errors	2.3	2.7	2.0	1.7	2.3	0.0
Workload	R Workload per Aircraft	0.2	0.3	0.3	0.3	0.2	0.1
	D Workload per Aircraft	0.2	0.2	0.2	0.2	0.2	0.0
	R Average Workload	2.0	4.7	5.3	4.7	3.0	1.0
	D Average Workload	2.0	3.7	3.0	3.3	3.0	0.0
	Communications per Aircraft	1.7	2.4	2.8	2.5	2.3	
	Data Entries per Aircraft	3.5	3.2	4.6	4.2	4.9	3.7

Note: All values are averaged across runs.

TABLE 9. SECTOR 27 - 12-MINUTE INTERVAL DATA

Construct	Variable	1	2	3	4	5	6	7
Safety	Data Block Offset and Leader Length	2.8	8.8	10.2	6.2	9.4	10.2	10.2
Capacity	Aircraft Under Control	8.0	14.0	13.0	8.2	8.6	16.0	13.2
	Altitude Assignments per Aircraft	0.5	0.5	0.5	0.2	0.5	0.9	0.7
Performance	R Data Entries	22.8	44.0	42.8	27.4	32.6	58.2	43.4
	R Data Entry Errors	0.8	3.0	1.8	1.0	2.8	2.0	1.2
	D Data Entries	10.4	10.2	5.2	2.4	47.8	15.0	5.2
	D Data Entry Errors	1.2	2.0	2.2	1.4	4.2	8.6	4.8
Workload	R Workload per Aircraft	0.3	0.3	0.3	0.2	0.2	0.3	0.4
	D Workload per Aircraft	0.2	0.2	0.2	0.2	0.2	0.3	0.3
	R Average Workload	2.0	3.8	3.6	1.8	1.6	4.8	5.0
	D Average Workload	1.6	2.2	2.4	1.4	1.4	4.0	4.0
	Communications per Aircraft	2.2	2.1	2.5	1.6	2.2	3.0	
	Data Entries per Aircraft	4.2	3.9	3.7	3.6	9.3	4.6	3.7

Note: All values are averaged across runs.

TABLE 10. SECTOR 35 - 12-MINUTE INTERVAL DATA

Construct	Variable	1	2	3	4	5	6	7
Safety	Data Block Offset and Leader Length	8.8	9.6	14.4	16.2	19.6	12.2	4.2
Capacity	Aircraft Under Control	9.0	12.8	16.4	22.4	27.0	21.6	10.8
	Altitude Assignments per Aircraft	0.0	0.0	0.1	0.1	0.1	0.1	0.1
Performance	R Data Entries	19.2	32.0	41.2	52.2	68.8	55.6	22.8
	R Data Entry Errors	0.4	1.2	2.6	2.0	3.8	1.6	0.4
	D Data Entries	12.6	7.2	5.2	16.2	47.2	5.0	5.6
	D Data Entry Errors	1.4	1.4	1.4	1.8	1.4	1.8	1.0
Workload	R Workload per Aircraft	0.2	0.2	0.2	0.2	0.2	0.2	0.2
	D Workload per Aircraft	0.2	0.2	0.2	0.2	0.1	0.2	0.3
	R Average Workload	1.8	2.4	3.0	4.4	4.6	3.6	2.2
	D Average Workload	1.8	2.6	2.6	3.4	1.8	3.6	3.6
	Communications per Aircraft	1.5	1.2	1.4	1.7	1.7	1.6	
	Data Entries per Aircraft	3.5	3.1	2.8	3.1	4.3	2.8	2.6

Note: All values are averaged across runs.

TABLE 11. SECTOR 38 - 12-MINUTE INTERVAL DATA

Construct	Variable	1	2	3	4	5	6
Safety	Data Block Offset and Leader Length	4.3	7.7	8.3	8.3	3.7	0.0
Capacity	Aircraft Under Control	10.0	12.7	13.3	11.3	11.0	5.0
	Altitude Assignments per Aircraft	0.5	0.3	0.3	0.4	0.5	0.4
Performance	R Data Entries	26.7	37.7	46.3	50.0	34.7	13.0
	R Data Entry Errors	1.0	0.7	2.7	4.0	3.3	1.0
	D Data Entries	11.7	6.3	7.0	5.0	16.3	1.0
	D Data Entry Errors	2.0	2.0	1.7	1.3	1.3	2.0
Workload	R Workload per Aircraft	0.2	0.2	0.2	0.1	0.1	0.2
	D Workload per Aircraft	0.1	0.2	0.2	0.1	0.2	0.6
	R Average Workload	2.0	2.0	2.3	1.7	1.3	1.0
	D Average Workload	1.0	2.3	2.0	1.7	1.7	3.0
	Communications per Aircraft	2.9	1.9	2.1	2.3	1.5	
	Data Entries per Aircraft	3.8	3.5	4.0	4.9	4.6	2.8

Note: All values are averaged across runs.

TABLE 12. SECTOR 26.- 12-MINUTE INTERVAL SD DATA

Construct	Variable	1	2	3	4	5
Safety	Data Block Offset and Leader Length	2.08	2.52	9.17	5.03	4.16
Capacity	Aircraft Under Control	0.58	0.58	2.08	1.15	1.00
	Altitude Assignments per Aircraft	1.00	1.00	4.51	3.46	4.16
Performance	R Data Entries	3.06	4.58	15.39	11.68	9.29
	R Data Entry Errors	1.53	2.89	1.73	0.00	1.15
	D Data Entries	3.00	4.04	8.19	4.16	5.57
	D Data Entry Errors	1.53	1.53	1.00	1.15	0.58
Workload	R Workload per Aircraft	0.11	0.12	0.06	0.17	0.21
	D Workload per Aircraft	0.07	0.06	0.18	0.12	0.08
	R Average Workload	1.00	2.08	1.53	2.08	3.00
	D Average Workload	0.58	1.15	3.00	1.53	1.00
	Communications per Aircraft	0.10	0.33	0.31	0.54	0.70
	Data Entries per Aircraft	0.11	0.12	0.06	0.17	0.21

Note: All values are averaged across runs.

TABLE 13. SECTOR 27 - 12-MINUTE INTERVAL SD DATA

Construct	Variable	1	2	3	4	5	6	7
Safety	Data Block Offset and Leader Length	1.30	4.92	3.96	5.40	7.50	5.12	4.44
Capacity	Aircraft Under Control	0.71	1.22	1.22	1.30	1.14	2.12	1.30
	Altitude Assignments per Aircraft	1.14	2.00	1.92	1.14	1.14	1.92	3.78
Performance	R Data Entries	2.95	2.55	5.17	5.08	6.99	6.98	8.17
	R Data Entry Errors	0.84	2.00	1.30	0.71	1.92	1.00	2.17
	D Data Entries	1.52	2.49	2.17	0.55	0.84	1.87	1.10
	D Data Entry Errors	0.45	1.00	1.30	0.89	1.48	5.37	5.22
Workload	R-Workload per Aircraft	0.09	0.12	0.08	0.11	0.11	0.06	0.09
	D Workload per Aircraft	0.14	0.08	0.11	0.07	0.13	0.15	0.16
	R Average Workload	0.71	1.48	0.89	0.84	0.89	1.30	1.58
	D Average Workload	0.89	1.10	1.52	0.55	1.14	2.00	1.87
	Communications per Aircraft	0.42	0.26	0.40	0.33	0.56	0.32	
	Data Entries per Aircraft	0.09	0.12	0.08	0.11	0.11	0.06	0.09

Note: All values are averaged across runs.

TABLE 14. SECTOR 35 - 12-MINUTE INTERVAL SD DATA

Construct	Variable	1	2	3	4	5	6	7
Safety	Data Block Offset and Leader Length	3.63	3.78	3.91	2.59	3.51	4.44	2.39
Capacity	Aircraft Under Control	0.71	0.84	1.67	5.00	1.00	1.34	1.79
	Altitude Assignments per Aircraft	0.55	0.45	1.41	0.45	0.71	0.45	0.45
Performance	R Data Entries	3.96	2.35	7.36	4.32	2.39	6.15	6.46
	R-Data Entry Errors	0.55	1.10	1.95	1.87	1.64	0.55	0.55
	D Data Entries	0.89	0.45	0.45	1.30	1.10	0.00	0.89
	D Data Entry Errors	1.67	1.52	1.52	2.17	1.52	1.79	1.00
Workload	R Workload per Aircraft	0.05	0.07	0.07	0.07	0.07	0.06	0.09
	D Workload per Aircraft	0.17	0.16	0.11	0.09	0.05	0.09	0.18
	R Average Workload	0.45	0.89	1.41	1.67	1.67	1.14	0.84
	D Average Workload	1.30	2.07	1.82	1.95	1.30	1.95	1.52
	Communications per Aircraft	0.41	0.27	0.27	0.26	0.25	0.21	
	Data Entries per Aircraft	0.05	0.07	0.07	0.07	0.07	0.06	0.09

Note: All values are averaged across runs.

TABLE 15. SECTOR 38 - 12-MINUTE INTERVAL SD DATA

Construct	Variable	1	2	3	4	5
Safety	Data Block Offset and Leader Length	0.58	3.06	3.06	2.31	3.79
Capacity	Aircraft Under Control	1.73	0.58	1.15	0.58	1.73
	Altitude Assignments per Aircraft	2.08	0.00	1.73	1.00	1.73
Performance	R Data Entries	3.06	1.15	7.51	1.73	10.02
	R Data Entry Errors	1.00	0.58	2.89	1.00	0.58
	D Data Entries	1.53	3.06	1.00	1.00	4.93
	D Data Entry Errors	1.73	1.73	1.53	1.53	2.31
Workload	R Workload per Aircraft	0.07	0.07	0.09	0.05	0.07
	D Workload per Aircraft	0.02	0.09	0.09	0.11	0.12
	R Average Workload	1.00	1.00	1.15	0.58	0.58
	D Average Workload	0.00	1.15	1.00	1.15	1.15
	Communications per Aircraft	0.23	0.13	0.37	0.37	0.73
	Data Entries per Aircraft	0.07	0.07	0.09	0.05	0.07

Note: All values are averaged across runs.

APPENDIX B
CONTROLLER COMMENTS

The following data represent controllers' partially-edited responses to the attached survey. Responses are organized by controller and section of the survey.

CONTROLLER A TEAM A

Section D

1. Radar and Map displays
Static Info
RAM limitations (data)

Need more flexibility in radar range selection, improved mapping (symbols, line types and color), ability to add geo-referenced lines and fix symbols dynamically (especially useful to support military missions). Static areas need to be improved to provide an electronic means of displaying maps, weather, approach charts, STARS/SIDS manuals, etc. At position current static information via paper cannot be located when needed! Need to boost available buffers for adaptation (pref routing and mapping currently affected most).

2. Misreading Flight progress strips
Making entries with keyboard
Equipment setup

Poor print quality of strips and general keyboard errors
Equipment setup - controller selection of each function required for operation (filter keys etc.)
- should have pre-set capability for intended operation (high or low sector configurations).

Section E

Primarily the lack of data (computer storage limitations and static data at position). Poor weather interface to PVD - this data should be available via electronic static displays above each operational position as well as other static data (charts, manuals, NOTAMS LOA's, directives, etc.).

CONTROLLER B TEAM A

Section D

1. Radar and Map displays
It would make things easier if the intensity of things on the map could be adjusted, i.e., have airspace boundaries brighter than airways.

2. Selecting targets with trackball
If you try to pull up a tag as the scope updates and the target moves you won't get the tag to display. The track ball should have bigger parameter to ID targets (sensitivity of trackball).

Section E

I think the equipment we use is fine. We have got more trouble with poor radar coverage and bad radios. To just update the PVD's and keep the old radar is not going to allow for increased service overall. The entire system needs to be changed not just the PVD.
This program is on the right track as long as it considers the needs of the controllers and the needs of the users. To just provide a new PVD without a change in procedures the same service will be provided.

CONTROLLER C TEAM B

Section D

1. Radar and Map displays - color maps with map filters would make it easier to identify airspace
2. Making entries with keyboard - mis-hitting function key without catching the mistake prior to entering command
-selecting targets with trackball - cursor must be in direct alignment with position symbol

Section E - blank

CONTROLLER D TEAM B

Section D

1. Radar and map displays
keyboard- layout is fine but the touch and sensitivity is deteriorating due to age, the keys stick
trackball - some are very stiff others move very quickly
volume of workspace - could use a little more room for handoff person
communications - R & L side handoff/override lines

-would like to be able to temporarily call up more features - range marks referenced to a fix to assist in spacing; all airports

2. Making entries with keyboard

Section E - blank

CONTROLLER E TEAM C

Section D

1. Flight strip bays - hard to keep reaching, hard to read, need to be electronically updated and marked
-Freq. switches - hard to find quickly - hard to reach

2. Misreading flight progress strips - They are set far from PVD dependent upon A side or D side to get strip there in a timely manner - red ink and black ink often malfunction - causes twice the work.
-making entries with keyboard - buttons stick, are too hard to find anyway

Section E - blank

CONTROLLER F TEAM D

Section D

1. none
2. none

Section E - blank

CONTROLLER G TEAM E

Section D

1. Trackball - you should be able to enter more data such as altitude with a simple entry

2. Selecting targets - you have to concentrate to make sure whatever you entered took

Section E - blank

CONTROLLER H TEAM E

Section D

1. Radar and map displays - could be improved in sharpness of display - windows tool box displays to adjust scope functions -tear off tool box displays which can be positioned at any position on the scope with pull down menus and hot buttons

Console Switches and knobs - VSCS should really help improve switches and knobs

Section E - blank

CONTROLLER I TEAM E

Section D

1. Other - The present system is highly effective. Of course having more data available at your disposal when you need it would enhance the control environment

2. Misreading Radar display information - constantly trying to increase working speed
Making entries with keyboard - working faster

Section E

If the up-time and reliability of the present system is maintained, the HCS/PVD/M1 system is adequate and we could continue with it for another few years until better technology is attained.

CONTROLLER J TEAM E

Section D

1. Radar and map displays -display needs better integration of weather

Flight strip bays -flight strips are unnecessary - redundant and should be automated/electronically - more a reference than a control tool

2. Misreading Radar display information
Misreading flight progress strips
-weather is usually erroneously displayed

Section E

An overall good system - a replacement tying the current functionality with new technology would be an ample update

CONTROLLER K TEAM F

Section D

1. Radar and Map displays - Radar needs color
Flight strip bays - no reference
Keyboard - is klunky and broken most of the time, it is fixed in position/not mobile - do not like lights that flash after entry
-need volume controls for all positions not one control for entire sector
-ambient noise of fans is loud

2. Making entries with keyboard - always sticks
Selecting targets with track ball - it is difficult to hit the position indicator with the trackball - the area for entry should be larger than the position indicator
adjusting the correct switch or knob - knobs all look the same

Section E - blank

CONTROLLER K TEAM F

Section D

1. Radar and Map displays - need the ability to request just the 1 area map you need instead of 2 or 3 areas on 1 map (MOA special use)
Keyboard - moveable keyboards - ability to touch type
other - noise - fan in this console is very noisy

2. Making entries with keyboard - eliminate need for most keyboard entries by allowing more entries with the trackball
other -hitting wrong QAK button- have the QAK buttons on screen- able to activate with the trackball

Section E

Teams should not have been assigned for all problems - you should have to work with other controllers to get different reactions to different combinations of people - we don't work the same way when working with some people as we do with others.

APPENDIX C
COMPARISONS TO FUTURE SYSTEMS

The purpose of this section is to provide guidance on using this baseline measurement methodology and the data on today's baseline system in making comparisons with future systems. The perspective taken is that quantitative baseline measurement data would be used in tandem with qualitative information for making assessments of future automation systems. The quantitative data are used to compute several statistics including an average or mean and a standard deviation[1] representing the degree of variability across controllers.

Qualitative information garnered from controllers and expert observers would be used to verify any issues or concerns identified through the analysis of the quantitative data. This information would be obtained during simulation run debriefings and a post-simulation caucus. This qualitative information would also be used to identify other issues or concerns not captured in the quantitative measurements but still pertinent in the comparison of a future system to today's baseline system.

The high level approach would consist of the following process:

a. Collect sufficient data on today's baseline system to provide stable estimates of all specified operational constructs and baseline measurements.

b. Reduce and analyze the data collected from "a" above, and complete the tables at each level of detail as shown in this report.

c. Collect the same data specified in "a" for the future system. Use the same airspace, simulation scenarios, controllers, and other aspects of the simulation that might otherwise work as intervening or confounding variables.

d. Complete the identical data reduction and analysis, specified in "b", for the future system.

e. Conduct a post-simulation caucus with the controllers and expert observers using the data comparisons from "b" and "d" as starting points to identify an initial set of issues and concerns. Use the data in other detailed tables to augment the analysis of these issues, as well as data contained in observer logs and debriefing materials. Make systematic comparisons between the PVD baseline and the future system, stepping through each quantitative measurement. Examine all data in a dynamic fashion to identify related trends that may or may not be found in other operational constructs and measurements to further substantiate or refute whether a problem exists.

During the caucus, consensus building techniques would be used with the controllers and observers to review and categorize the quantitative comparisons, facilitate closure in identifying and prioritizing significant issues, and assess the viability of potential operational resolutions

[1] Standard deviation is the average difference of controllers' scores from the mean score.

(e.g., changes to operational procedures, shifts in training duration, and technique). This may necessitate participation of procedures and training specialists in these assessments.

As part of the assessment, it will be necessary to verify that a problem is not an artifact of the simulation platform, a functional or performance deficiency of the operational software, or some other intervening variable potentially skewing the comparisons between the two systems.

An important basis for determining whether the future system is comparable to the baseline system is whether the data for any particular measure are statistically equivalent. That is, do the two systems numerically share the same average or have overlapping ranges or confidence intervals. However, statistical equivalence or non-equivalence does not automatically indicate operational equivalence or non-equivalence. This must be determined by expert judgement. Results of these comparisons could be categorized as follows:

a. Category 1 involves measurements where the baseline and future systems have data that are statistically equivalent and are deemed to be operationally equivalent.

b. Category 2 involves measurements where the baseline and future systems have data that are statistically equivalent but are deemed to be operationally different.

c. Category 3 involves measurements where the baseline and future systems have data that are not statistically equivalent but the systems are deemed to be operationally equivalent.

d. Category 4 involves measurements where the baseline and future systems have data that are not statistically equivalent and the systems are deemed to be operationally different.

Statistical equivalency is determined on the basis of traditional descriptive and inferential statistics. A preliminary approach to use of these statistics is as follows:

a. Descriptive statistics making general comparisons of means, standard deviations, and trends. These comparisons are appropriate for data contained in appendix A, tables 1, 4 and 5.

b. Inferential statistics, such as using an analysis of variance (ANOVA) with post-hoc testing to compare the baseline and future systems on a given measurement. ANOVAs are pertinent to appendix A, tables 2, 6, and 8 through 11. Those ANOVAs would be configured as two-way ANOVAs comprised of two factors:

1. Systems (i.e., PVD baseline versus the future system).
2. A second factor consisting of one of:

    a) the four sectors in table 2.
    b) the four sectors in table 6.

c) the two operational positions (R and D) in table 6.
d) the time segments in tables 8 through 11.

The ANOVA first checks for a difference in one of the above main factors (1 and 2) and then for an interaction between the two factors. If statistically significant differences are found, then post-hoc testing would identify where the difference(s) occur.

In addition, a t-test or non-parametric test would be appropriate for table 4 on a selective basis. This test would evaluate differences in NAS-message use between the PVD baseline and a future system.

The following is an example demonstrating the use of an ANOVA. An alpha level (or margin for error) should be adapted based upon an operational projection of the power of the test. It is assumed, pending verification when further simulation data are obtained, that ATC measurements are normally distributed, permitting use of parametric statistics. Non-parametric statistics may be appropriate for some measures. Statistical tests can be used as a technique to compare systems, but do not eliminate the need for a controller caucus.

As an example of using statistics, consider the baseline measurement of the average workload for the radar controller. Appendix A, table 2 contains the means for this measure across the four sectors on today's PVD baseline system. These means, along with hypothetical means for a future system, are shown below.

	Sector 26	Sector 27	Sector 35	Sector 38
PVD	2.8	2.4	2.8	2.5
Future System	5.5	5.5	3.5	3.5

These means are graphically depicted in figure 1.

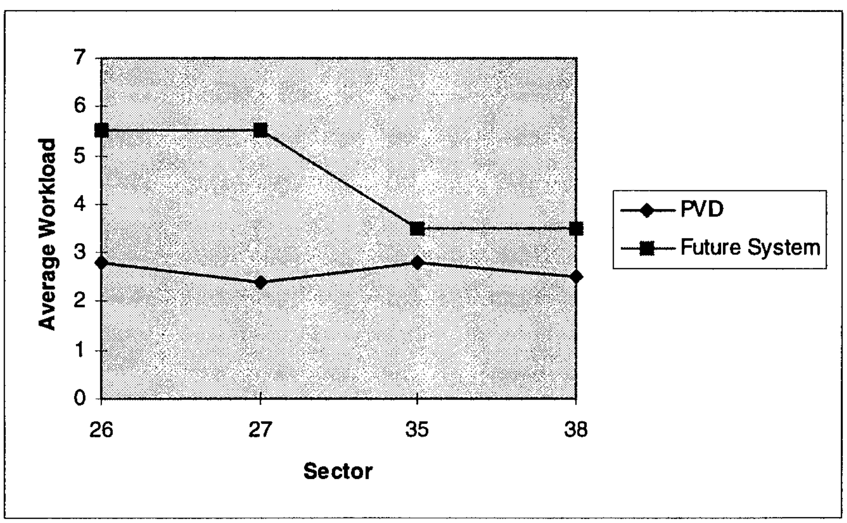


FIGURE 1. COMPARISON OF PVD AND FUTURE SYSTEMS

The ANOVA would test for an overall difference between the PVD and a future system and for differences between sectors. It would also test the statistical significance of the interaction represented in the above figure. The presence of an interaction means that there is a differential effect in how a measure (such as workload) changes across the two variables (systems and sectors). If the ANOVA shows significant overall effects or a significant interaction, then post-hoc testing would be done to determine where the difference(s) occur. This post-hoc testing might show the hypothetical future system has significantly greater workload than the PVD system for the low altitude sectors. Even though the future system might show somewhat higher workload values for the high altitude sectors, the difference may not reach statistical significance.

Computational techniques for ANOVA are readily available in a large number of statistics books and commercial software programs.

APPENDIX D
FURTHER PVD BASELINING SIMULATION

Planning for further PVD baselining simulation is being conducted to provide a more stable set of data. A detailed report defining plans and specifying evaluation test procedures will be developed prior to conducting the simulation. This planning document will identify the time to accommodate training, data collection, and controller debriefings. Assignments of controllers from the en route user team to data collection sectors and DYSIM positions will be specified. The document will also contain materials for airspace and procedures training. Questionnaires, logs, and other data collection materials will be included. For continuity and comparability with the baselining data contained in this report and to meet AT requirements, the same ZDC airspace and the simulation scenarios will be used.

Controllers coming from different en route facilities will be trained using pre-meeting mailings of materials, classroom training, and hands-on familiarization. Packages containing materials on ZDC airspace, sector procedures, and traffic flow information will be mailed prior to the meeting. Classroom training will address sector airspace, traffic flows, local procedures, overall study methodology, the post-simulation run questionnaire and debrief, the Air Traffic Workload Input Technique (ATWIT) workload evaluation methodology, final questionnaire, the role of expert observers, and other factors. Sufficient hands-on training of the controllers on this unfamiliar airspace will be provided to achieve a threshold of proficiency.

☆U.S. GOVERNMENT PRINTING OFFICE: 1996 - 704-066/20195